



## SEQUENCE LISTING

*10346X*  
<110> Mayfield, Stephen

<120> RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR  
EXPRESSION OF RECOMBINANT MOLECULES

<130> SCR2177S

<140> 09/341, 550  
<141> 1999-07-13

<150> PCT/US98/00840  
<151> 1998-01-16

<150> 60/035, 955  
<151> 1997-01-17

<150> 60/069, 400  
<151> 1997-12-12

<160> 14

<170> PatentIn Ver. 2.1

<210> 1  
<211> 15  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 1  
Gln Tyr Gly Phe Val His Phe Glu Asp Gln Ala Ala Ala Asp Arg  
1 5 10 15

*B*  
<210> 2  
<211> 14  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 2  
Gly Phe Gly Phe Ile Asn Phe Lys Asp Ala Glu Ser Ala Ala  
1 5 10

<210> 3  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
oligonucleotide primer

<400> 3  
cagtacgggt tcgtbcaytt cgaggaycag gc

32

<210> 4

<211> 40  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
oligonucleotide primer

<400> 4 40  
ggaattcggg ttcgggttca tyaacttcaa ggaygcbgag

<210> 5  
<211> 2846  
<212> DNA  
<213> Chlamydomonas reinhardtii

<220>  
<221> CDS  
<222> (197)..(2065)

<400> 5 60  
gaattcgcgg ccgcgtccgtg gttgggtcctc atgggtgtctt tttgaagagg acctgagcct 120  
ttcacccaaa tatatcaaaa aaccggggca accggccaaa aaaattgcaa aagcctctcg 180  
taggcacaaa agaccttattc tagccatcaa ctttgtatcc gacgctgccg ttttagctg 232  
cgtcttgaag tcaagg atg gcg act act gag tcc tcg gcc ccg gcg gcc acc  
Met Ala Thr Thr Glu Ser Ser Ala Pro Ala Ala Thr  
1 5 10

acc cag ccg gcc agc acc ccg ctg gcg aac tcg tcg ctg tac gtc ggt 280  
Thr Gln Pro Ala Ser Thr Pro Leu Ala Asn Ser Ser Leu Tyr Val Gly  
15 20 25

30 35 40  
gac ctg gag aag gat gtc acc gag gcc cag ctg ttc gag ctc ttc tcc 328  
Asp Leu Glu Lys Asp Val Thr Glu Ala Gln Leu Phe Glu Leu Phe Ser

45 50 55 60  
tcg gtt ggc cct gtg gcc tcc att cgc gtg tgc cgc gat gcc gtc acg 376  
Ser Val Gly Pro Val Ala Ser Ile Arg Val Cys Arg Asp Ala Val Thr

65 70 75  
cgc cgc tcg ctg ggc tac gcc tac gtc aac tac aac agc gct ctg gac 424  
Arg Arg Ser Leu Gly Tyr Ala Tyr Val Asn Tyr Asn Ser Ala Leu Asp

80 85 90  
ccc cag gct gct gac cgc gcc atg gag acc ctg aac tac cat gtc gtg 472  
Pro Gln Ala Ala Asp Arg Ala Met Glu Thr Leu Asn Tyr His Val Val

95 100 105  
aac ggc aag cct atg cgc atc atg tgg tcg cac cgc gac cct tcg gcc 520  
Asn Gly Lys Pro Met Arg Ile Met Trp Ser His Arg Asp Pro Ser Ala

110 115 120  
cgc aag tcg ggc gtc ggc aac atc ttc atc aag aac ctg gac aag acc 568  
Arg Lys Ser Gly Val Gly Asn Ile Phe Ile Lys Asn Leu Asp Lys Thr

atc gac gcc aag gcc ctg cac gac acc ttc tcg gcc ttc ggc aag att		616
Ile Asp Ala Lys Ala Leu His Asp Thr Phe Ser Ala Phe Gly Lys Ile		
125	130	135
140		
ctg tcc tgc aag gtt gcc act gac gcc aac ggc gtg tcg aag ggc tac		664
Leu Ser Cys Lys Val Ala Thr Asp Ala Asn Gly Val Ser Lys Gly Tyr		
145	150	155
ggc ttc gtg cac ttc gag gac cag gcc gct gcc gat cgc gcc att cag		712
Gly Phe Val His Phe Glu Asp Gln Ala Ala Asp Arg Ala Ile Gln		
160	165	170
acc gtc aac cag aag aag att gag ggc aag atc gtg tac gtg gcc ccc		760
Thr Val Asn Gln Lys Lys Ile Glu Gly Lys Ile Val Tyr Val Ala Pro		
175	180	185
ttc cag aag cgc gct gac cgc ccc agg gca agg acg ttg tac acc aac		808
Phe Gln Lys Arg Ala Asp Arg Pro Arg Ala Arg Thr Leu Tyr Thr Asn		
190	195	200
gtg ttc gtc aag aac ttg ccg gcc gac atc ggc gac gac gag ctg ggc		856
Val Phe Val Lys Asn Leu Pro Ala Asp Ile Gly Asp Asp Glu Leu Gly		
205	210	215
220		
aag atg gcc acc gag cac ggc gag atc acc acg gcg gtg gtc atg aag		904
Lys Met Ala Thr Glu His Gly Glu Ile Thr Ser Ala Val Val Met Lys		
225	230	235
gac gac aag ggc ggc agc aag ggc ttc ggc ttc atc aac ttc aag gac		952
Asp Asp Lys Gly Gly Ser Lys Gly Phe Gly Phe Ile Asn Phe Lys Asp		
240	245	250
gcc gag tcg gcg gcc aag tgc gtg gag tac ctg aac gag cgc gag atg		1000
Ala Glu Ser Ala Ala Lys Cys Val Glu Tyr Leu Asn Glu Arg Glu Met		
255	260	265
270		
agc ggc aag acc ctg tac gcc ggc cgc gcc cag aag aag acc gag cgc		1048
Ser Gly Lys Thr Leu Tyr Ala Gly Arg Ala Gln Lys Lys Thr Glu Arg		
275	280	
gag gcg atg ctg cgc cag aag gcc gag gag agc aag cag gag cgt tac		1096
Glu Ala Met Leu Arg Gln Lys Ala Glu Glu Ser Lys Gln Glu Arg Tyr		
285	290	295
300		
ctg aag tac cag agc atg aac ctg tac gtc aag aac ctg tcc gac gag		1144
Leu Lys Tyr Gln Ser Met Asn Leu Tyr Val Lys Asn Leu Ser Asp Glu		
305	310	315
320		
gag gtc gac gac gac gcc ctg cgt gag ctg ttc gcc aac tct ggc acc		1192
Glu Val Asp Asp Asp Ala Leu Arg Glu Leu Phe Ala Asn Ser Gly Thr		
325	330	
335		
atc acc tcg tgc aag gtc atg aag gac ggc agc ggc aag tcc aag ggc		1240
Ile Thr Ser Cys Lys Val Met Lys Asp Gly Ser Gly Lys Ser Lys Gly		
340	345	
350		
ttc ggc ttc gtg tgc ttc acc agc cac gac gag gcc acc cgg ccc ccc		1288
Phe Gly Phe Val Cys Phe Thr Ser His Asp Glu Ala Thr Arg Pro Pro		
355	360	

gtg acc gag atg aac ggc aag atg gtc aag ggc aag ccc ctg tac gtg	1336
Val Thr Glu Met Asn Gly Lys Met Val Lys Gly Lys Pro Leu Tyr Val	
365 370 375 380	
gcc ctg gcg cag cgc aag gac gtg cgc cgt gcc acc cag ctg gag gcc	1384
Ala Leu Ala Gln Arg Lys Asp Val Arg Arg Ala Thr Gln Leu Glu Ala	
385 390 395	
aac atg cag gcg cgc atg ggc atg ggc gcc atg agc cgc ccg ccg aac	1432
Asn Met Gln Ala Arg Met Gly Met Gly Ala Met Ser Arg Pro Pro Asn	
400 405 410	
ccg atg gcc ggc atg agc ccc tac ccc ggc gcc atg ccg ttc ttc gct	1480
Pro Met Ala Gly Met Ser Pro Tyr Pro Gly Ala Met Pro Phe Phe Ala	
415 420 425	
ccc ggc ccc ggc ggc atg gct gct ggc ccg cgc gct ccg ggc atg atg	1528
Pro Gly Pro Gly Gly Met Ala Ala Gly Pro Arg Ala Pro Gly Met Met	
430 435 440	
tac ccg ccc atg atg ccg ccg cgc ggc atg cct ggc ccc ggc cgc ggc	1576
Tyr Pro Pro Met Met Pro Pro Arg Gly Met Pro Gly Pro Gly Arg Gly	
445 450 455 460	
ccc cgc ggc ccc atg atg ccg ccc cag atg atg ggt ggc ccc atg atg	1624
Pro Arg Gly Pro Met Met Pro Pro Gln Met Met Gly Gly Pro Met Met	
465 470 475	
ggc ccg ccc atg ggc ccc ggg cgc ggc cgt ggc ggc cgc ggc ccc tcc	1672
Gly Pro Pro Met Gly Pro Gly Arg Gly Arg Gly Gly Arg Gly Pro Ser	
480 485 490	
ggc cgc ggc cag ggc cgc ggc aac aac gcc cct gcc cag cag ccc aag	1720
Gly Arg Gly Gln Gly Arg Gly Asn Asn Ala Pro Ala Gln Gln Pro Lys	
495 500 505	
ccc gcc gct gag ccg gcc gcc gcg ccc gcc gcc gcc ccc gct gcc	1768
Pro Ala Ala Glu Pro Ala Ala Pro Ala Ala Ala Pro Ala Ala	
510 515 520	
gcg gcg cct gcc gcc gcg gag ccg gag gcc ccc gcc gcc cag cag	1816
Ala Ala Pro Ala Ala Ala Glu Pro Glu Ala Pro Ala Ala Gln Gln	
525 530 535 540	
ccg ctg acc gcc tcc gcg ctg gcc gcc gcg ccg gag cag cag aag	1864
Pro Leu Thr Ala Ser Ala Leu Ala Ala Ala Pro Glu Gln Gln Lys	
545 550 555	
atg atg atc ggc gag cgc ctg tac ccg cag gtg gcg gag ctg cag ccc	1912
Met Met Ile Gly Glu Arg Leu Tyr Pro Gln Val Ala Glu Leu Gln Pro	
560 565 570	
gac ctg gct ggc aag atc acc ggc atg ctg ctg gag atg gac aac gcc	1960
Asp Leu Ala Gly Lys Ile Thr Gly Met Leu Leu Glu Met Asp Asn Ala	
575 580 585	
gag ctt ctg atg ctt ctg gag tcg cac gag gcg ctg gtg tcc aag gtg	2008
Glu Leu Leu Met Leu Leu Glu Ser His Glu Ala Leu Val Ser Lys Val	
590 595 600	

gac gag gcc atc gct gtg ctc aag cag cac aac gtg att gcc gag gag 2056  
Asp Glu Ala Ile Ala Val Leu Lys Gln His Asn Val Ile Ala Glu Glu  
605 610 615 620

aac aag gct taaagcgct gcacgcttgt gcgggctggc ggccgcggc 2105  
Asn Lys Ala

cgccgcggc ctgcttggc cgccggcagc atggcgccgg cggacgcggc gtgggagcag 2165  
tgcttgcgtc ttctggccgc cgtgaagccg cgccgaactg gggcggacgg caggctggc 2225  
ttgacgcccgg cgccgcacaa cacaaagttt gtggcgtgaa agtctctgg cgtgctccgg 2285  
acggttgtaa ggttttaaga actggctttt ggccgggtt ccgcggaaag gcggAACGGC 2345  
ggtctttca ggcacatcac atccggctgg aaaaattctt accaaagcca acccctgcac 2405  
ccaaaaattt cgggttccga aagaacactc ccctttttc cggcaacgcg ttcttcaag 2465  
gcacatcact ttccgggtt gaaagaaaatg ttacccggaa aaggcggaa gccccctgca 2525  
cccgacaaag ttattcgggg ttccggccgg aatgagcaag cgttcgggct gttggccgta 2585  
tcgcacgc tgtcgggtt ctaggcgcca gaaggaagga tgacgtttt gtgaagggtt 2645  
gcaaaactgag cacacgagtt ttggcaatag acgtggagaa agtccagtgc ggggtgaggc 2705  
ggatagcggg atcaagcgtg gcgggtccct ggcgagacga gacgcttctg ttgttttgc 2765  
gagccctttt atggcacaat cgcactgttt tgagcaggcg actgtaaagt gcccgcgc 2825  
aaaaaaagcgg ccgcgaattc c 2846

<210> 6  
<211> 15  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 6  
Trp Phe Val Asp Gly Glu Leu Ala Ser Asp Tyr Asn Gly Pro Arg  
1 5 10 15

<210> 7  
<211> 22  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 7  
Gln Leu Ile Leu Trp Thr Thr Ala Asp Asp Leu Lys Ala Asp Ala Glu  
1 5 10 15  
Ile Met Thr Val Phe Arg  
20

<210> 8

```

<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      oligonucleotide primer

<400> 8
cgcgatccg aygcbgagat yatgac 26

<210> 9
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      oligonucleotide primer

<400> 9
cgcgaattcg tcatratctc vgcrtc 26

<210> 10
<211> 2413
<212> DNA
<213> Chlamydomonas reinhardtii

<220>
<221> CDS
<222> (16)..(1614)

<400> 10
gagtagttt acgcc atg aac cgt tgg aac ctt ctt gcc ctt acc ctg ggg 51
      Met Asn Arg Trp Asn Leu Leu Ala Leu Thr Leu Gly
      1           5           10

ctg ctg ctg gca gcg ccc ttc acc aag cac cag ttt gct cat gct 99
Leu Leu Leu Val Ala Ala Pro Phe Thr Lys His Gln Phe Ala His Ala
      15          20          25

tcc gat gag tat gag gac gac gag gag gac gat gcc ccc gcc gcc cct 147
Ser Asp Glu Tyr Glu Asp Asp Glu Glu Asp Asp Ala Pro Ala Ala Pro
      30          35          40

aag gac gac gac gtc gac gtt act gtg gtg acc gtc aag aac tgg gat 195
Lys Asp Asp Asp Val Asp Val Thr Val Val Thr Val Lys Asn Trp Asp
      45          50          55          60

gag acc gtc aag aag tcc aag ttc gcg ctt gtg gag ttc tac gct cct 243
Glu Thr Val Lys Ser Lys Phe Ala Leu Val Glu Phe Tyr Ala Pro
      65          70          75

tgg tgc ggc cac tgc aag acc ctc aag cct gag tac gct aag gct gcc 291
Trp Cys Gly His Cys Lys Thr Leu Lys Pro Glu Tyr Ala Lys Ala Ala
      80          85          90

```

acc gcc ctg aag gct gct ccc gat gcc ctt atc gcc aag gtc gac	339
Thr Ala Leu Lys Ala Ala Ala Pro Asp Ala Leu Ile Ala Lys Val Asp	
95 100 105	
gcc acc cag gag gag tcc ctg gcc cag aag ttc ggc gtg cag ggc tac	387
Ala Thr Gln Glu Glu Ser Leu Ala Gln Lys Phe Gly Val Gln Gly Tyr	
110 115 120	
ccc acc ctc aag tgg ttc gtt gat ggc gag ctg gct tct gac tac aac	435
Pro Thr Leu Lys Trp Phe Val Asp Gly Glu Leu Ala Ser Asp Tyr Asn	
125 130 135 140	
ggc ccc cgc gac gct gat ggc att gtt ggc tgg gtg aag aag aag act	483
Gly Pro Arg Asp Ala Asp Gly Ile Val Gly Trp Val Lys Lys Lys Thr	
145 150 155	
ggc ccc ccc gcc gtg acc gtt gag gac gcc gac aag ctg aag tcc ctg	531
Gly Pro Pro Ala Val Thr Val Glu Asp Ala Asp Lys Leu Lys Ser Leu	
160 165 170	
gag gcg gac gct gag gtc gtt gtc gtc ggc tac ttc aag gcc ctg gag	579
Glu Ala Asp Ala Glu Val Val Val Gly Tyr Phe Lys Ala Leu Glu	
175 180 185	
ggc gag atc tac gac acc ttc aag tcc tac gcc gcc aag acc gag gac	627
Gly Glu Ile Tyr Asp Thr Phe Lys Ser Tyr Ala Ala Lys Thr Glu Asp	
190 195 200	
gtg gtg ttc gtg cag acc acc agc gcc gac gtc gcc aag gcc ggc	675
Val Val Phe Val Gln Thr Thr Ser Ala Asp Val Ala Lys Ala Ala Gly	
205 210 215 220	
ctg gac gcc gtg gac acc gtg tcc gtg gtc aag aac ttc gcc ggt gag	723
Leu Asp Ala Val Asp Thr Val Ser Val Val Lys Asn Phe Ala Gly Glu	
225 230 235	
gac cgc gcc acc gcc gtc ctg gcc acg gac atc gac act gac tcc ctg	771
Asp Arg Ala Thr Ala Val Leu Ala Thr Asp Ile Asp Thr Asp Ser Leu	
240 245 250	
acc gcg ttc gtc aag tcg gag aag atg ccc ccc acc att gag ttc aac	819
Thr Ala Phe Val Lys Ser Glu Lys Met Pro Pro Thr Ile Glu Phe Asn	
255 260 265	
cag aag aac tct gac aag atc ttc aac agc ggc atc aac aag cag ctg	867
Gln Lys Asn Ser Asp Lys Ile Phe Asn Ser Gly Ile Asn Lys Gln Leu	
270 275 280	
att ctg tgg acc acc gcc gac gac ctg aag gcc gac gcc gag atc atg	915
Ile Leu Trp Thr Thr Ala Asp Asp Leu Lys Ala Asp Ala Glu Ile Met	
285 290 295 300	
act gtg ttc cgc gag gcc agc aag aag ttc aag ggc cag ctg gtg ttc	963
Thr Val Phe Arg Glu Ala Ser Lys Lys Phe Lys Gly Gln Leu Val Phe	
305 310 315	
gtg acc gtc aac aac gag ggc gac ggc gcc gac ccc gtc acc aac ttc	1011
Val Thr Val Asn Asn Glu Gly Asp Gly Ala Asp Pro Val Thr Asn Phe	
320 325 330	

ttc ggc ctc aag ggc gcc acc tcg cct gtg ctg ctg ggc ttc ttc atg Phe Gly Leu Lys Gly Ala Thr Ser Pro Val Leu Leu Gly Phe Phe Met 335	340	345	1059	
gag aag aac aag aag ttc cgc atg gag ggc gag ttc acg gct gac aac Glu Lys Asn Lys Lys Phe Arg Met Glu Gly Glu Phe Thr Ala Asp Asn 350	355	360	1107	
gtg gct aag ttc gcc gag agc gtg gtg gac ggc acc gcg cag gcc gtg Val Ala Lys Phe Ala Glu Ser Val Val Asp Gly Thr Ala Gln Ala Val 365	370	375	380	1155
ctc aag tcg gag gcc atc ccc gag gac ccc tat gag gat ggc gtc tac Leu Lys Ser Glu Ala Ile Pro Glu Asp Pro Tyr Glu Asp Gly Val Tyr 385	390	395	1203	
aag att gtg ggc aag acc gtg gag tct gtg gtt ctg gac gag acc aag Lys Ile Val Gly Lys Thr Val Glu Ser Val Val Leu Asp Glu Thr Lys 400	405	410	1251	
gac gtg ctg ctg gag gtg tac gcc ccc tgg tgc ggc cac tgc aag aag Asp Val Leu Leu Glu Val Tyr Ala Pro Trp Cys Gly His Cys Lys Lys 415	420	425	1299	
ctg gag ccc atc tac aag aag ctg gcc aag cgc ttt aag aag gtg gat Leu Glu Pro Ile Tyr Lys Leu Ala Lys Arg Phe Lys Lys Val Asp 430	435	440	1347	
tcc gtc atc atc gcc aag atg gat ggc act gag aac gag cac ccc gag Ser Val Ile Ile Ala Lys Met Asp Gly Thr Glu Asn Glu His Pro Glu 445	450	455	460	1395
atc gag gtc aag ggc ttc cct acc atc ctg ttc tat ccc gcc ggc agc Ile Glu Val Lys Gly Phe Pro Thr Ile Leu Phe Tyr Pro Ala Gly Ser 465	470	475	1443	
gac cgc acc ccc atc gtg ttc gag ggc ggc gac cgc tcg ctc aag tcc Asp Arg Thr Pro Ile Val Phe Glu Gly Gly Asp Arg Ser Leu Lys Ser 480	485	490	1491	
ctg acc aag ttc atc aag acc aac gcc aag atc ccg tac gag ctg ccc Leu Thr Lys Phe Ile Lys Thr Asn Ala Lys Ile Pro Tyr Glu Leu Pro 495	500	505	1539	
aag aag ggc tcc gac ggc gac gag ggc acc tcg gac gac aag gac aag Lys Lys Gly Ser Asp Gly Asp Glu Gly Thr Ser Asp Asp Lys Asp Lys 510	515	520	1587	
ccc gcg tcc gac aag gac gag ctg taa gcggctatct gaactacccc Pro Ala Ser Asp Lys Asp Glu Leu 525	530		1634	
aggtttggag cgtctgcttg cgcgcttgcg cttgcacact gtgcacatggat gggagtttaag 1694				
gaggagacgg agcacggagg ctgcgctcggttgggtggctt ggagcacccgg cagcgcgtga 1754				
tccgtcctgg cagcagcaac ggccggagccgg ggcgcatttg ggcgcgagctg ggcgcggct 1814				
gttgctggag aggatatgct gcccggccgg aggaaggct agggcagag atgagagcgt 1874				

tacgggctgg catgcgggcg cccgtgcctc tccctgcggc gcagtccttgc ctaggagacg 1934  
cacggttttg ccaaagaggg acgctgtcca cagccctgcgc actggaaagtt ttttaggccc 1994  
tgccgtggta gtgggtttgg tacggttgtg tgcataagat gaacaacgtt tctctcaaga 2054  
cgagactact agtatgctga cgggtgtgt atgtggtgga tggattgtgc cccgaccatg 2114  
aagagtgctg tggtgcctcg gcgcattctgt cgcctggat gtgcgtggtt ccgaacgctg 2174  
gagtcatctg ttgaggagcg aggggtttgt cgggtccgccc cggcacggcc gcgtgatgtc 2234  
cgatgggga ttgcgagcga gggcaaccgc agcgcagata gcgcgcgcgc ggatcgagct 2294  
agcgcaggat gatgagagcc gggcattcgc ggcgtggat cagggaggag ccaaggcgg 2354  
gtgcatgcga ggaaaacagt gtgcggcaa gaacgggctg caagaacgcc ttgcgcaa 2413

<210> 11  
<211> 4  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 11  
Cys Gly His Cys  
1

<210> 12  
<211> 4  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 12  
Lys Asp Glu Leu  
1

<210> 13  
<211> 1424  
<212> DNA  
<213> Chlamydomonas reinhardtii

<220>  
<221> CDS  
<222> (252)..(1310)

<220>  
<221> misc\_feature  
<222> (279)  
<223> Codon also can encode Ser

<220>  
<221> misc\_feature  
<222> (282)  
<223> Codon also can encode Glu

<220>

<221> misc\_feature  
<222> (294)  
<223> Codon also can encode Gly

<220>  
<221> misc\_feature  
<222> (306)  
<223> Codon also can encode Asn

<220>  
<221> misc\_feature  
<222> (357)  
<223> Codon also can encode Leu

<220>  
<221> misc\_feature  
<222> (369)  
<223> Codon also can encode Thr

<220>  
<221> misc\_feature  
<222> (486)  
<223> Codon also can encode Ser

<220>  
<221> misc\_feature  
<222> (495)  
<223> Codon also can encode Ile

<220>  
<221> misc\_feature  
<222> (510)  
<223> Codon also can encode Ala

<220>  
<221> misc\_feature  
<222> (555)  
<223> Codon also can encode Val

<220>  
<221> misc\_feature  
<222> (588)  
<223> Codon also can encode Glu

<220>  
<221> misc\_feature  
<222> (600)  
<223> Codon also can encode Leu

<220>  
<221> misc\_feature  
<222> (621)  
<223> Codon also can encode Ala

<220>  
<221> misc\_feature  
<222> (714)  
<223> Codon also can encode Thr

```
<220>
<221> misc_feature
<222> (729)
<223> Codon also can encode Ile

<220>
<221> misc_feature
<222> (1092)
<223> Codon also can encode Val

<220>
<221> misc_feature
<222> (1191)
<223> Codon also can encode Ile

<220>
<221> misc_feature
<222> (1284)
<223> Codon also can encode Ala

<220>
<221> misc_feature
<222> (1287)
<223> Codon also can encode Ile

<220>
<221> misc_feature
<222> (1290)
<223> Codon also can encode Glu

<220>
<221> misc_feature
<222> (1293)
<223> Codon also can encode Ala

<220>
<221> misc_feature
<222> (1296)
<223> Codon also can encode Pro

<220>
<221> misc_feature
<222> (1302)
<223> Codon also can encode Thr

<220>
<221> misc_feature
<222> (1308)
<223> Codon also can encode Gly

<400> 13
cgtcctattt taatactccg aaggaggcag ttggcaggca actgccactg acgtcccgta 60
aggtaaggg gacgtccact ggcgtcccgta aaggggaaagg ggacgttaggt acataaatgt 120
gctaggtaac taacgttga tttttgtgg tataatatat gtaccatgct tttaatagaa 180
gcttgaattt ataaataaa atattttac aatatttac ggagaaattha aaactttaaa 240
```

aaaat	taaca	t	atg	aca	gca	att	tta	gaa	cgt	cgt	gaa	aat	tct	agc	cta	290
			Met	Thr	Ala	Ile	Leu	Glu	Arg	Arg	Glu	Asn	Ser	Ser	Leu	
			1				5					10				
tgg	gct	cgt	ttt	tgt	gag	tgg	atc	act	tca	act	gaa	aac	cgt	tta	tac	338
Trp	Ala	Arg	Phe	Cys	Glu	Trp	Ile	Thr	Ser	Thr	Glu	Asn	Arg	Leu	Tyr	
			15				20					25				
atc	ggt	tgg	ttc	ggt	gta	atc	atg	atc	cca	tgt	ctt	ctt	act	gca	aca	386
Ile	Gly	Trp	Phe	Gly	Val	Ile	Met	Ile	Pro	Cys	Leu	Leu	Thr	Ala	Thr	
	30				35					40			45			
tca	gta	ttc	atc	atc	gct	ttc	atc	gct	gct	ccg	cca	gta	gac	atc	gat	434
Ser	Val	Phe	Ile	Ile	Ala	Phe	Ile	Ala	Ala	Pro	Pro	Val	Asp	Ile	Asp	
			50				55					60				
ggt	atc	cgt	gaa	cca	gta	ttc	atc	ggt	tct	ctt	tac	ggt	aac	aac	atc	482
Gly	Ile	Arg	Glu	Pro	Val	Ser	Gly	Ser	Leu	Leu	Tyr	Gly	Asn	Asn	Ile	
	65					70					75					
att	aca	ggt	gct	gta	atc	cca	act	tct	aac	gca	atc	ggt	ctt	cac	ttc	530
Ile	Thr	Gly	Ala	Val	Ile	Pro	Thr	Ser	Asn	Ala	Ile	Gly	Leu	His	Phe	
	80				85					90						
tac	cca	att	tgg	gaa	gct	gct	tct	cta	gac	gag	tgg	tta	tac	aac	ggt	578
Tyr	Pro	Ile	Trp	Glu	Ala	Ala	Ser	Leu	Asp	Glu	Trp	Leu	Tyr	Asn	Gly	
	95				100					105						
ggt	cct	tac	caa	ctt	atc	ggt	tgt	cac	ttc	ctt	cta	ggt	gta	tac	tgc	626
Gly	Pro	Tyr	Gln	Leu	Ile	Val	Cys	His	Phe	Leu	Leu	Gly	Val	Tyr	Cys	
	110				115					120			125			
tac	atg	ggt	cgt	gag	tgg	gaa	tta	tct	ttc	cgt	tta	ggt	atg	cgt	cca	674
Tyr	Met	Gly	Arg	Glu	Trp	Glu	Leu	Ser	Phe	Arg	Leu	Gly	Met	Arg	Pro	
	130					135					140					
tgg	atc	gct	gta	gct	tac	tca	gct	cca	gta	gct	gca	gct	tca	gct	gta	722
Trp	Ile	Ala	Val	Ala	Tyr	Ser	Ala	Pro	Val	Ala	Ala	Ala	Ser	Ala	Val	
	145				150					155						
ttc	tta	gtt	tac	cct	atc	ggc	caa	ggt	tca	ttc	tct	gac	ggt	atg	cct	770
Phe	Leu	Val	Tyr	Pro	Ile	Gly	Gln	Gly	Ser	Phe	Ser	Asp	Gly	Met	Pro	
	160					165					170					
tta	ggt	atc	tct	ggt	act	ttc	aac	ttc	atc	gta	ttc	caa	gca	gaa	818	
Leu	Gly	Ile	Ser	Gly	Thr	Phe	Asn	Phe	Met	Ile	Val	Phe	Gln	Ala	Glu	
	175					180					185					
cac	aac	atc	ctt	atg	cac	cca	ttc	cac	atg	tta	ggt	ggt	gct	ggt	gta	866
His	Asn	Ile	Leu	Met	His	Pro	Phe	His	Met	Leu	Gly	Val	Ala	Gly	Val	
	190				195					200			205			
ttc	ggt	ggt	tca	tta	ttc	tca	gct	atg	cac	ggt	tct	tta	gtt	act	tca	914
Phe	Gly	Gly	Ser	Leu	Phe	Ser	Ala	Met	His	Gly	Ser	Leu	Val	Thr	Ser	
	210					215					220					
tct	tta	atc	cgt	gaa	aca	act	gaa	aac	gaa	tca	gct	aac	gaa	ggt	tac	962
Ser	Leu	Ile	Arg	Glu	Thr	Thr	Glu	Asn	Glu	Ser	Ala	Asn	Glu	Gly	Tyr	
	225					230					235					

cgt ttc ggt caa gaa gaa gaa act tac aac att gta gct gct cat ggt	1010
Arg Phe Gly Gln Glu Glu Glu Thr Tyr Asn Ile Val Ala Ala His Gly	
240 245 250	
tac ttt ggt cgt cta atc ttc caa tac gct tct ttc aac aac tct cgt	1058
Tyr Phe Gly Arg Leu Ile Phe Gln Tyr Ala Ser Phe Asn Asn Ser Arg	
255 260 265	
tca tta cac ttc ttc tta gct gct tgg ccg gta atc ggt att tgg ttc	1106
Ser Leu His Phe Leu Ala Ala Trp Pro Val Ile Gly Ile Trp Phe	
270 275 280 285	
act gct tta ggt tta tca act atg gca ttc aac tta aac ggt ttc aac	1154
Thr Ala Leu Gly Leu Ser Thr Met Ala Phe Asn Leu Asn Gly Phe Asn	
290 295 300	
ttc aac caa tca gta gta gac tca caa ggt cgt gta cta aac act tgg	1202
Phe Asn Gln Ser Val Val Asp Ser Gln Gly Arg Val Leu Asn Thr Trp	
305 310 315	
gca gac atc atc aac cgt gct aac tta ggt atg gaa gta atg cac gag	1250
Ala Asp Ile Ile Asn Arg Ala Asn Leu Gly Met Glu Val Met His Glu	
320 325 330	
cgt aac gct cac aac ttc cct cta gac tta gct tca act aac tct agc	1298
Arg Asn Ala His Asn Phe Pro Leu Asp Leu Ala Ser Thr Asn Ser Ser	
335 340 345	
tca aac aac taa tttttttta aactaaaata aatctggta accataccta	1350
Ser Asn Asn	
350	
gtttatttta gtttatacac acttttcata tatataact taatagctac cataggcagt	1410
tggcaggacg tccc	1424

<210> 14  
 <211> 1278  
 <212> DNA  
 <213> Chlamydomonas reinhardtii

<220>  
 <221> CDS  
 <222> (1)..(1272)

atg ggc cat cat cat cat cat cat cat cac agc agc ggc cat	48
Met Gly His His His His His His His His Ser Ser Gly His	
1 5 10 15	
atc gaa ggt cgt cat atg gcg act act gag tcc tcg gcc ccg gcg gcc	96
Ile Glu Gly Arg His Met Ala Thr Thr Glu Ser Ser Ala Pro Ala Ala	
20 25 30	
acc acc cag ccg gcc agc acc ccg ctg gcg aac tcg tcg ctg tac gtc	144
Thr Thr Gln Pro Ala Ser Thr Pro Leu Ala Asn Ser Ser Leu Tyr Val	
35 40 45	

ggt gac ctg gag aag gat gtc acc gag gcc cag ctg ttc gag ctc ttc	192
Gly Asp Leu Glu Lys Asp Val Thr Glu Ala Gln Leu Phe Glu Leu Phe	
50 55 60	
tcc tcg gtt ggc cct gtg gcc tcc att cgc gtg tgc cgc gat gcc gtc	240
Ser Ser Val Gly Pro Val Ala Ser Ile Arg Val Cys Arg Asp Ala Val	
65 70 75 80	
acg cgc cgc tcg ctg ggc tac gcc tac gtc aac tac aac agc gct ctg	288
Thr Arg Arg Ser Leu Gly Tyr Ala Tyr Val Asn Tyr Asn Ser Ala Leu	
85 90 95	
gac ccc cag gct gct gac cgc gcc atg gag acc ctg aac tac cat gtc	336
Asp Pro Gln Ala Ala Asp Arg Ala Met Glu Thr Leu Asn Tyr His Val	
100 105 110	
gtg aac ggc aag cct atg cgc atc atg tgg tcg cac cgc gac cct tcg	384
Val Asn Gly Lys Pro Met Arg Ile Met Trp Ser His Arg Asp Pro Ser	
115 120 125	
gcc cgc aag tcg ggc ggc aac atc ttc atc aag aac ctg gac aag	432
Ala Arg Lys Ser Gly Val Gly Asn Ile Phe Ile Lys Asn Leu Asp Lys	
130 135 140	
acc atc gac gcc aag gcc ctg cac gac acc ttc tcg gcc ttc ggc aag	480
Thr Ile Asp Ala Lys Ala Leu His Asp Thr Phe Ser Ala Phe Gly Lys	
145 150 155 160	
att ctg tcc tgc aag gtt gcc act gac gcc aac ggc gtg tcg aag ggc	528
Ile Leu Ser Cys Lys Val Ala Thr Asp Ala Asn Gly Val Ser Lys Gly	
165 170 175	
tac ggc ttc gtg cac ttc gag gac cag gcc gct gcc gat cgc gcc att	576
Tyr Gly Phe Val His Phe Glu Asp Gln Ala Ala Asp Arg Ala Ile	
180 185 190	
cag acc gtc aac cag aag aag att gag ggc aag atc gtg tac gtg gcc	624
Gln Thr Val Asn Gln Lys Lys Ile Glu Gly Lys Ile Val Tyr Val Ala	
195 200 205	
ccc ttc cag aag cgc gct gac cgc ccc agg gca agg acg ttg tac acc	672
Pro Phe Gln Lys Arg Ala Asp Arg Pro Arg Ala Arg Thr Leu Tyr Thr	
210 215 220	
aac gtg ttc gtc aag aac ttg ccg gcc gac atc ggc gac gac gag ctg	720
Asn Val Phe Val Lys Asn Leu Pro Ala Asp Ile Gly Asp Asp Glu Leu	
225 230 235 240	
ggc aag atg gcc acc gag cac ggc gag atc acc agc gcg gtg gtc atg	768
Gly Lys Met Ala Thr Glu His Gly Glu Ile Thr Ser Ala Val Val Met	
245 250 255	
aag gac gac aag ggc ggc agc aag ggc ttc ggc ttc atc aac ttc aag	816
Lys Asp Asp Lys Gly Gly Ser Lys Gly Phe Gly Ile Asn Phe Lys	
260 265 270	
gac gcc gag tcg gcg gcc aag tgc gtg gag tac ctg aac gag cgc gag	864
Asp Ala Glu Ser Ala Ala Lys Cys Val Glu Tyr Leu Asn Glu Arg Glu	
275 280 285	

atg	agc	ggc	aag	acc	ctg	tac	gcc	ggc	cgc	gcc	cag	aag	aag	acc	gag	912
Met	Ser	Gly	Lys	Thr	Leu	Tyr	Ala	Gly	Arg	Ala	Gln	Lys	Lys	Thr	Glu	
290					295						300					
cgc	gag	gag	atg	ctg	cgc	cag	aag	gcc	gag	gag	agc	aag	cag	gag	cgt	960
Arg	Glu	Ala	Met	Leu	Arg	Gln	Lys	Ala	Glu	Glu	Ser	Lys	Gln	Glu	Arg	
305				310					315						320	
tac	ctg	aag	tac	cag	agc	atg	aac	ctg	tac	gtc	aag	aac	ctg	tcc	gac	1008
Tyr	Leu	Lys	Tyr	Gln	Ser	Met	Asn	Leu	Tyr	Val	Lys	Asn	Leu	Ser	Asp	
325						330					335					
gag	gag	gtc	gac	gac	gcc	ctg	cgt	gag	ctg	ttc	gcc	aac	tct	ggc	1056	
Glu	Glu	Val	Asp	Asp	Asp	Ala	Leu	Arg	Glu	Leu	Phe	Ala	Asn	Ser	Gly	
340						345					350					
acc	atc	acc	tcg	tgc	aag	gtc	atg	aag	gac	ggc	agc	ggc	aag	tcc	aag	1104
Thr	Ile	Thr	Ser	Cys	Lys	Val	Met	Lys	Asp	Gly	Ser	Gly	Lys	Ser	Lys	
355						360					365					
ggc	ttc	ggc	ttc	gtg	tgc	ttc	acc	agc	cac	gac	gag	gcc	acc	cgg	ccg	1152
Gly	Phe	Gly	Phe	Val	Cys	Phe	Thr	Ser	His	Asp	Glu	Ala	Thr	Arg	Pro	
370						375					380					
ccc	gtg	acc	gag	atg	aac	ggc	aag	atg	gtc	aag	ggc	aag	ccc	ctg	tac	1200
Pro	Val	Thr	Glu	Met	Asn	Gly	Lys	Met	Val	Lys	Gly	Lys	Pro	Leu	Tyr	
385						390					395				400	
gtg	gcc	ctg	gag	cag	cgc	aag	gac	gtg	cgc	cgt	gcc	acc	cag	ctg	gag	1248
Val	Ala	Leu	Ala	Gln	Arg	Lys	Asp	Val	Arg	Arg	Ala	Thr	Gln	Leu	Glu	
							405		410			415				
gcc	aac	atg	cag	gag	ccg	atg	taa	ggatcc								1278
Ala	Asn	Met	Gln	Ala	Arg	Met										
							420									

B1  
Concluded